PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 10 AUGUST 15, 1924 Number 8

ON THE SIMULTANEOUS JUMPING OF TWO ELECTRONS IN BOHR'S MODEL

By Paul S. Epstein

Normal Bridge Laboratory of Physics, Pasadena

Communicated, June 11, 1924

1. Messrs. H. N. Russell and F. A. Saunders recently discovered in the arc spectrum of calcium a new series[1] ($1p-mp'$) which they attribute to the simultaneous jumping of two electrons. It must be conceded that the evidence in favor of such a hypothesis is very strong. There cannot be the slightest doubt that the lines in question really constitute a series, because Russell and Saunders have divided them into groups showing the regular triplet structure, each group containing the same spacings (105.9 and 52.3) due to the three levels of the first term. This term is, therefore, identified as ($1p$) of the triplet system ($1p_1 = 33988.7$, $1p_2 = 34094.6$, $1p_3 = 34146.9$).[2] Beginning with the third group of the series the second terms then turn out to be negative.[3] This means that the energy available for emission is larger than that corresponding to a jump of an electron from infinity to the $1p$-level. It is hard to see how the excess of energy could be supplied, unless it is due to the jumping of a second electron. Moreover, such a conclusion is strongly supported by quantitative relations. If we consider the limiting value of the second term corresponding to the limit of the series or to the removal of the first electron into infinity, this whole value must be due to the energy of the second electron. At the same time the atom has now the constitution of an ionized atom, so that the above limiting value must correspond to one of the possible energy differences in that state of the atom or, in other words, to one of the frequencies of the spark spectrum. This is in fact the case: as Russell and Saunders pointed out, the extrapolated limiting value falls within 1/200 volt from the energy difference $1\delta-1\pi$ in the spark spectrum, or within the limits of the error of extrapolation. We have, therefore, to regard the occurrence of such double jumps as an established experimental fact and to look for a theoretical explanation of it.

2. At first sight it seems that it is not easy to reconcile the conditions

sketched in the preceding section with the principle of correspondence. However, it will be shown that a closer study of the problem leads to the conclusion that there is no difficulty: on the contrary, the phenomenon in question, far from contradicting the principle of correspondence, flows from it as its natural consequence and is *due to the coupling of the electrons within the atom.* To make this clear we shall use a very much simplified model of an atom containing two electrons which, however, retains all the essential features of a real atom with regard to the question we are going to study.

3. Let our model be constituted of two electrons attached by forces proportional to the distance to two positions of rest O_1 and O_2 and able to move only in the straight line O_1O_2. Let the distance between O_1 and O_2 be denoted by a and the displacements of the electrons from their respective positions of rest by x_1 and x_2. We assume that these displacements are small compared with a, and that a is small compared with the wave length of the emitted radiation. The distance between the two electrons is, therefore, equal to $a + x_2 - x_1$ and the force of their mutual interaction to $e^2/(a + x_2 - x_1)^2$, if e means the charge of an electron. Denoting the coefficients of the elastic forces of attachment to the two positions of rest by $-m\omega_1^2$ and $-m\omega_2^2$ (m being the mass of the electron), we finally get the equations of motion in the form

$$\ddot{x}_1 + \omega_1^2 x_1 = -e^2/m(a + x_2 - x_1)^2, \qquad \ddot{x}_2 + \omega_2^2 x_2 = e^2/m(a + x_2 - x_1)^2 \quad (1)$$

4. We are going to discuss first the case when the distance a is so large that the force of coupling is negligible compared to the forces of attachment. Equations (1) assume then the simple form

$$\ddot{x}_1 + \omega_1^2 x_1 = 0, \qquad \ddot{x}_2 + \omega_2^2 x_2 = 0 \quad (2)$$

defining the harmonic oscillations

$$x_1 = A_1 \cos(\omega_1 t + \delta_1), \qquad x_2 = A_2 \cos(\omega_2 t + \delta_2) \quad (3)$$

Our system represents the totality of two linear resonators. Subjecting the motion of these resonators to quantum restrictions, we obtain for the energy of the stationary orbits the well known expression of Planck[4] $n_1\omega_1 h/2\pi$ for the first electron and $n_2\omega_2 h/2\pi$ for the second, where n_1 and n_2 represent quantic integers. The total energy of the system is, therefore

$$E = h(n_1\omega_1 + n_2\omega_2)/2\pi \quad (4)$$

According to Bohr's theory light is emitted by our system if its motion changes from the stationary state characterized by the numbers n_1', n_2' to the state corresponding to n_1, n_2, the frequency ν_Q being determined by the difference of energies divided by

$$\nu_Q = [(n_1' - n_1)\omega_1 + (n_2' - n_2)\omega_2]/2\pi \quad (5)$$

5. The question is now, which of the possible differences $n_1' - n_1$, $n_2' - n_2$ do actually occur. And this question is answered by Bohr's principle of correspondence, which is based on the study of the emission of light which the same system would have in the classical theory. The amplitude of the emitted waves is according to the classical theory proportional to the total acceleration of electric charges taking place in the system, that is, to $\ddot{x}_1 + \ddot{x}_2$ (provided that the size of the system is small compared with the wave length). For a system of two degrees of freedom with a double periodicity characterized by the two angular velocities ω_1 and ω_2 the general expression of the acceleration is the two-fold Fourier series

$$\ddot{x}_1 + \ddot{x}_2 = \sum_0^\infty \sum_0^\infty A_{\tau_1\tau_2} \cos [(\tau_1\omega_1 + \tau_2\omega_2)t + \delta_{\tau_1\tau_2}] \qquad (6)$$

where τ_1 and τ_2 denote whole numbers. The possible frequencies are, therefore, given in the classical theory by the arguments of all the terms of the Fourier series

$$\nu_c = (\tau_1\omega_1 + \tau_2\omega_2)/2\pi \qquad (7)$$

We see that expressions (5) and (7) become identical if we put

$$n_1' - n_1 = \tau_1, \qquad n_2' - n_2 = \tau_2 \qquad (8)$$

so that to every possible frequency in the quantum theory there corresponds a possible classical frequency. However, in the classical case a large part of the frequencies contained in formula (7) will not actually occur, because many of the coefficients $A_{\tau_1\tau_2}$ will turn out to be equal to zero. For instance, if we compute $\ddot{x}_1 + \ddot{x}_2$ for our model, we have from (3)

$$\ddot{x}_1 + \ddot{x}_2 = -\omega_1^2 A_1 \cos(\omega_1 t + \delta_1) - \omega_2^2 A_2 \cos(\omega_2 t + \delta_2) \qquad (9)$$

so that the double sum (6) is reduced to only two terms and our of the possible cases (7) only those two will be realized in which $\tau_1 = 1$, $\tau_2 = 0$, and $\tau_1 = 0$, $\tau_2 = 1$.

The principle of correspondence can be briefly stated as the hypothesis that light of any frequency ν_Q in the theory of quanta has the same intensity as the corresponding frequency ν_c computed on the basis of the classical theory. It follows from this and relations (8) that out of the totality of possible jumps represented by formula (5) only two will actually occur, namely

$$1.\ n_1' - n_1 = 1,\ n_2' - n_2 = 0; \qquad 2.\ n_1' - n_1 = 0,\ n_2' - n_2 = 1 \qquad (10)$$

These relations give the *selection principle* for our model: the quantic numbers can change only by zero or by unity, and only one of the electrons can jump at a time.

6. Turning to the general case, when the coupling cannot be neglected, we make use of our assumption that x_1 and x_2 are small compared with a. We can then expand the right side of equations (1) in powers of $(x_2-x_1)/a$

$$\frac{e^2}{m(x_2-x_1+a)^2} = \frac{e^2}{ma^2} - \frac{2e^2}{ma^3}(x_2-x_1) + \frac{3e^2}{ma^4}(x_2-x_1)^2 + \ldots$$

The change in the motion of our electrons which the two first terms of this series will produce is well known. The constant term produces only a displacement of the positions of rest, the linear term a slight change in the frequency of the oscillations. Neither of these effects is of any interest to us. We may, therefore, treat the problem as if these two terms were absent and restrict ourselves to the discussion of the third quadratic term. With the abbreviation $3e^2/ma^4 = 2s$ equations (1) assume the form

$$\ddot{x}_1 + \omega_1^2 x_1 = -2s(x_2-x_1)^2; \qquad \ddot{x}_2 + \omega_2^2 x_2 = 2s(x_2-x_1)^2 \qquad (11)$$

The term on the right side is a small correction term. We get, therefore, the first approximation if we substitute into this term the approximate values (3) for x_1 and x_2:

$$\begin{aligned}\ddot{x}_1 + \omega_1^2 x_1 = {} & -s(A_1^2 + A_2^2) + 2sA_1A_2 \cos[(\omega_1+\omega_2)t + \delta_1 + \delta_2] + \\ & + 2sA_1A_2 \cos[(\omega_1-\omega_2)t + \delta_1 - \delta_2] + sA_1^2 \cos(2\omega_1 t + 2\delta_1) + \\ & + sA_2^2 \cos(2\omega_2 t + 2\delta_2)\end{aligned} \qquad (12)$$

Owing to the inhomogeneous terms on the right, we must, in order to get a solution of this equation add to the integral (9) the particular integral of this equation which is composed just of the same cosine terms that occur in the equation only with different coefficients. In this way we find for the acceleration

$$\begin{aligned}\ddot{x}_1 + \ddot{x}_2 = {} & -\omega_1^2 A_1 \cos(\omega_1 t + \delta_1) - \omega_2^2 A_2 \cos(\omega_2 t + \delta_2) + \\ & + A_{11} \cos(2\omega_1 t + \delta_{11}) + A_{12} \cos[(\omega_1+\omega_2)t + \delta_{12}] + \\ & + A_{21} \cos[(\omega_1-\omega_2)t + \delta_{21}] + A_{22} \cos(2\omega_2 t + \delta_{22})\end{aligned} \qquad (13)$$

We have written the coefficients of the additional terms in an abbreviated form instead of giving their expressions in terms of A_1, A_2, ω_1, ω_2. It is sufficient for our purpose to notice that they do not vanish and that they are, in general, all of the same order of magnitude. With our new equations (11) the energy expression (4) will no longer be strictly valid. We are, however, interested not in the exact values of the resulting frequencies, but in the permissible combinations of quantic numbers. The correction, is, therefore, of no avail for our purpose and we may continue to use formula (5) in its old shape. The same reasoning as in section 5 leads to expect, on account of the principle of correspondence, the following additional combinations:

3. $n_1'-n_1 = 2, \quad n_2'-n_2 = 0,$ 4. $n_1'-n_1 = 1, \quad n_2'-n_2 = 1,$
5. $n_1'-n_1 = \pm 1, \quad n_2'-n_2 = \mp 1,$ 6. $n_1'-n_1 = 0, \quad n_2'-n_2 = 2.$

We see that the selection principle (10) breaks down in two respects. Besides the change of a quantic number by *one* or *zero* there appears a change by *two* units. Moreover, both quantic numbers may change by unity at the same time, and *these last combinations correspond to the simultaneous jumping of both electrons*. It is obvious that by discussing the cubic term of our expansion and taking a model with three electrons we would obtain the possibility of three electrons jumping at the same time. However, the corresponding lines must have a much lower intensity.

6. If we regard instead of our simplified model the actual conditions in an atom, the results will not change essentially. The selection principle as stated by Bohr and Rubinowicz is that the azimutal quantum number k for a jumping electron can change only by unity: $k'-k = \pm 1$. If we neglect in an atom containing several electrons their mutual interactions we will, as in section 4, get the additional restriction that only one electron can jump at a time. On the other hand, the coupling of the electrons will cause this selection principle to break down and new combinations, including the simultaneous jumping of two electrons, to become permissible. In fact, it is not so hard to calculate the perturbations in question as to present the calculations in a concise form. The first new combinations to appear are $k'-k = \pm 2$, $k'-k = 0$ and the simultaneous jumping of two electrons ($k'-k = \pm 1$ for each of them). The combination $k'-k = 0$ has a simple geometrical meaning. The absence of it is due to the orbit being a plane one. When the coupling forces become so strong that the motion of the electron can no longer be regarded as plane even for the short time of its period, the combination $k'-k = 0$ assumes an appreciable probability.

Which of the three new combinations appears first with increasing coupling depends on the size, shape and relative orientation of the two orbits and can hardly be predicted in a general way. It is, however, interesting to note that the double jump discovered by Russell and Saunders appears in the spectra of metals of the second group where the Bohr-Rubinowicz selection principle just begins to break down, and the combinations $k'-k = \pm 2$ and $k'-k = 0$ begin to become a regular feature of the spectrum, especially in the singlet system. It may be expected, therefore, that in the spectra of the groups of higher order, where the selection principle breaks down completely, jumps of this type will become much more common.

[1] H. N. Russell and F. A. Saunders (to be published).

[2] The first two groups of this series were found by R. Götze (*Ann. Physik, Leipzig*, **66**, p. 291, 1921.)

[3] Negative terms occur also in another series of the calcium spectrum $1d-md'$. The first group was given by Götze (*loc. cit.*). Russell and Saunders succeeded in finding the second group in which the terms $2d'$ are negative. It is most probable that the $1p-1p'$ series of strontium is produced in quite the same way as the analogous series of calcium, but only the first two groups of the strontium series are known.

[4] As $\omega_1/2\pi = \nu_1$, the expression may be written $n_1 h\nu_1$. Compare: M. Planck, *Wärmestrahlung*, section 137. Leipzig, 1921.

THE COMPTON AND DUANE EFFECTS

By Joseph A. Becker[1]

Norman Bridge Laboratory of Physics, Pasadena

Communicated June 19, 1924

Late in 1922, Professor A. H. Compton[2] developed a quantum theory for the scattering of X-rays by collisions with electrons. According to this theory when radiation of wave-length λ falls on a scattering substance the scattered wave-length increases by an amount given by the formula

$$\Delta\lambda = \frac{2h}{mc}\sin^2\frac{\theta}{2} = 0.0484\sin^2\frac{\theta}{2} \qquad (1)$$

where θ is the angle between the direction of the incident and emergent beams. We shall call this the Compton shift.

Compton, using an ionization spectrometer with low dispersion, actually found the wave-length shifted by the predicted amount. P. A. Ross[3] at Stanford University repeated the experiment using a photographic plate to detect the radiation and fairly high resolving power. He not only found the correct Compton shift but also an unshifted line, which is most likely due to collisions of the radiation with massive nuclear systems. On repeating his experiment, Compton also found this unshifted line which in some cases is more intense than the shifted one.

Bergen Davis at Columbia University also reported finding these two lines but according to him the Compton shift depends slightly on the atomic number of the scattering substance.

Since then the Compton shift has been verified also in England in the Davy-Faraday Laboratory by Muller. He used the K α-rays of silver, molybdenum, and copper scattered by paraffin, glass, and aluminum. All together these workers have verified the Compton shift as given by formula (1) for incident wave-lengths ranging from 0.2 to 1.5 Å. with scattering substances well spread over the periodic table and for various angles.

On the other hand Professor William Duane[4] and his collaborators working with the ionization method have been unable to find any evidence for the Compton shift. They do, however, find a new line or band whose short wave-length limit is shifted toward longer wave-lengths by an amount given by

$$\Delta\lambda = \frac{\lambda_1^2}{\lambda_2 - \lambda_1} \qquad (2)$$

where λ_1 = incident λ and λ_2 = critical absorption λ of the scattering element. Their theory of tertiary radiation accounts for this Duane shift. They have used tungsten K radiation with carbon, aluminum, sulphur, copper, molybdenum, and silver scatterers as well as molybdenum K radiation falling on lithium, carbon, aluminum, sulphur, sodium chloride, and ice.

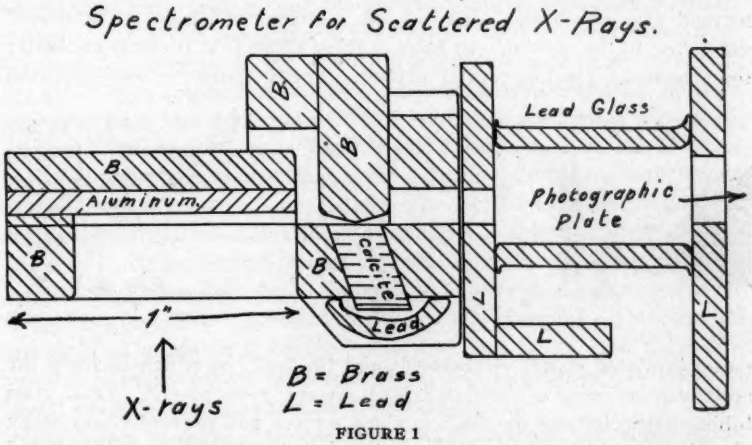


FIGURE 1

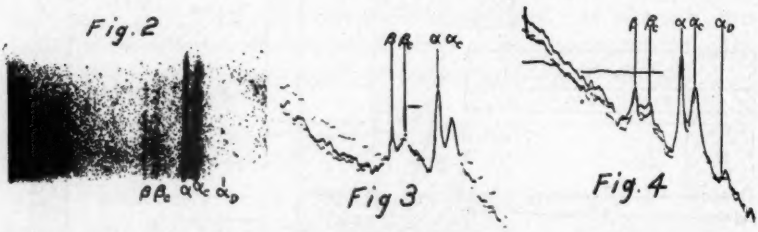
In all these cases they find evidence for their tertiary radiation which varies with the scattering substance but in no case do they see any evidence for the Compton shift.

There can be little doubt that these two effects are distinct and that both exist. What remains to be explained is why Duane does not find the Compton effect and vice-versa. Of course there are differences in the experimental arrangements but at first sight none of these seem essential.

To test out the suggestion first proposed by Mr. Smythe of this laboratory that the essential difference between the methods was one of intensity, the following experiments were performed. A standard water cooled molybdenum target Coolidge tube was used throughout. It was operated at about 40 kilovolts peak without a rectifier. In all the experiments aluminum was used as scatterer and the angle between the incident and scattered beam was about 100°. Under these conditions the

Compton shift should amount to 0.027Å. while the Duane tertiary radiation shift should be 0.069Å. for α_1. The separation α_1 to β for molybdenum amounts to 0.077Å. and for α_1 to α_2 it is 0.004Å. In other words the Compton shift amounts to about six times the α doublet separation or to about one-third of the $\alpha_1\beta$ separation, while the Duane shift for α_1 should be about nine-tenths of the $\alpha_1\beta$ separation. With aluminum it should therefore be readily possible to get both Compton and Duane shifts on the same plate for both α and β lines.

Figure 1 shows in cross-section the relative positions and dimensions of the X-ray tube, the scatterer, and the calcite crystal used as a Seemann spectrograph. It will be noticed that the aluminum subtends a large enough angle at the slit such that all of the wave-lengths which are of interest can be reflected at all times without moving the crystal. This method also allows the source of the scattered radiation to be moved very close to the crystal. In both of these respects it is distinctly better than the Bragg rotating crystal method. The spectrometer was supported



on an insulated stand. Between it and the lead box which enclosed the apparatus a series of insulated lead screens were placed. Lead glass tubing connected the openings in these screens and prevented any stray direct or scattered radiation from getting into the passage from the crystal to the photographic plate. It was found essential to screen the aluminum, crystal, and photographic plate from all possible stray radiations. The spectrograph gave clear sharp lines whose widths were only slightly greater than twice the slit opening and did not increase when the distance from the crystal was increased four-fold.

Film 2 was placed 42 cm. from the crystal and exposed for 61 hours with 25 ma. through the tube. While the impressions are too faint for reproduction they show a distinct Compton shift for both α and β lines. The α doublet is clearly resolved into its two components and the lines are quite sharp. The magnitude of the shift is 0.0275Å. for each of the three lines. There is no indication of a Duane shift.

Since we were primarily interested in studying how these shifts changed as the intensity falling on the aluminum was decreased rather than in the dispersion, the photographic film for the following exposures was placed

21 cm. from the crystal. The dispersion was now such that the $\alpha_1\alpha_2$ lines just overlapped while the $\alpha_1\beta$ separation was 2.56 mm. The width of the lines was about 0.2 mm.

Film 3 was exposed for 61 hours. It shows the unshifted α doublet and β line together with the Compton shifted lines. The latter were slightly less intense and only a trifle wider than the unshifted lines. The size of the shift is about 0.026Å. for α and 0.025Å. for β. There is no indication of a Duane shift for either α or β.

Film 4 was a repetition of 3 in all details except that the current through the tube was 8.5 ma. for 180 hours. In other words, the intensity is one-third while the time is multiplied by three. This film shows not only the shifted and unshifted α and β lines with about the same relative intensity as in 3, but also a faint new line, α_D, whose short wave-length edge is displaced from the α peak by just the amount predicted by Duane's tertiary radiation. This film is reproduced in Figure 2. α_D is too weak to show up. Through the courtesy of Mr. Pettit films 3 and 4 were photometered at the Mt. Wilson observatory. Three independent runs over three different regions of each film were made. These were then juxtaposed and photographed. They are shown in Figure 3 and 4, respectively. The fact that the lower intensity film does show a slight Duane shift is some confirmation of the hypothesis that the essential difference between Compton's and Duane's experimental procedure is one of intensity. If in our experiment the intensity could be reduced further by a factor of three it might very well be that the Duane shift would become much more pronounced compared to the Compton shift. To follow up this clue would require a whole months exposure. The author is unable to continue this work. The suggestion together with its partial experimental support is published in the hope that it will help clear the situation in the controversy.

The author is indebted to Professor Ross of Stanford University for his keen interest and valuable suggestions in regard to experimental technique. He is also indebted to Professor Watson and Messrs. Brode, Mott-Smith, and Du-Mond of the local institute for their suggestions and help in setting up the apparatus.

NOTE (July 25). At Stanford University Professors Webster, Ross and I measured with an ionization chamber the total intensities of Radiation from blocks of graphite aluminum and lead illuminated by rays from a molybdenum target seven centimeters away. The tube was run at 40 kilovolts with currents of 1.25, 2.5, 6.25, 12.5, and 25 milliamperes. This total radiation should include both Compton's and Duane's rays and could not be expected to be proportional to the primary, if the above theory of intensity effects is correct, except by an extraordinary coincidence. The fact is however that it is proportional to better than 1%. D. L. WEBSTER, J. A. BECKER.

[1] NATIONAL RESEARCH FELLOW.

[2] Compton, *Bull. Nat. Res. Coun., Washington*, Oct., 1922. See also *Physic. Rev.*,

Ithaca, Feb. 1923, p. 207; May 1923, p. 483; June 1923, p. 715; *Phil. Mag. London*, Nov. 1923, p. 897.

[3] Ross, these PROCEEDINGS, July 1923, and June 1924.

[4] Duane and others, these PROCEEDINGS, Dec. 1923, p. 419; Jan. 1924, p. 41; March 1924, p. 92; April 1924, p. 148.

COULOMB'S LAW AND THE HYDROGEN SPECTRUM

BY EDWIN B. WILSON

HARVARD SCHOOL OF PUBLIC HEALTH, BOSTON

Communicated, July 3, 1924

In the Bohr theory of the hydrogen atom restricted, let us say, to circular orbits so as to deal only with the simplest case, Coulomb's law has been assumed to hold. This is carrying over to the microcosm our molar laws of electricity and might seem a doubtful procedure in the face of our giving up so much of our mechanics just as it seems to some questionable, after taking the electron as the indivisible electric unit, to talk of the distribution of electricity over or through the electron.[1] In favor of Coulomb's law for the microcosm we have our general tendency to carry over and apply old laws whenever they work and insofar as they work. Further in favor are the scattering experiments of Rutherford dealing with the positive nuclei.[2] And it may be that best of all our evidence is the success with which theories of spectra have been worked out by combining Coulomb's law with elementary mechanics plus the quantum hypothesis.

Is then the law of Coulomb really implied by the quantum theory? The equations of motion for the electron in its circle are

$mrw^2 = -F = dV/dr$	force equation	(1)
$mr^2w = nh/2\pi$	quantum condition	(2)
$E_2 - E_1 = h\nu$	frequency condition	(3)
$\nu = N(1/n_1^2 - 1/n_2^2)$	spectral law	(4)

Here F is the central force, V is the potential energy, r is the orbital radius, m is the mass, w is the angular velocity, E is the total energy, etc.

$$E = V + \frac{1}{2}mr^2w^2 \qquad \text{energy equation} \qquad (5)$$

From (4) and (3) we have

$$E_1 + Nh/n_1^2 = E_2 + Nh/n_2^2 \qquad (6)$$

For all subscript combinations arising from $n_1 = 1,2,\ldots$, $n_2 = 2,3,\ldots$, $n_1 < n_2$. Hence

$$V_i + \frac{1}{2} mr_i^2 w_i^2 + Nh/n_i^2 = \text{constant}, \; i = 1,2,\ldots \qquad (7)$$

As V may contain an additive constant we may write

$$V + \frac{1}{2} mr^2 w^2 + Nh/n^2 = 0 \qquad (8)$$

This equation must hold for all values of r which occur in the orbits. Let us assume it holds identically for all values of r, as it surely holds for all those in which we are interested and as the use of $-F = dV/dr$ implies other values of r at least in the vicinity of the orbital radii.

Now equations (1), (2), (8) contain in addition to r and V the unknown or variable quantities w and n. We shall eliminate them to obtain a relation between V and r. First introduce the auxiliary variable $u = 1/r^2$. Then

$$\frac{1}{2} mw^2 = -u^2 dV/du \qquad \text{from (1)}$$

$$2\pi mw = unh \qquad \text{from (2)}$$

$$V + \frac{1}{2} mw^2/u + Nh/n^2 = 0 \qquad \text{from (8)}$$

Hence $V - udV/du = -Nh/n^2$ from (1) and (8)

But $$-\frac{1}{n^2} = -\frac{u^2h^2}{4\pi^2m^2w^2} = \frac{h^2}{8\pi^2 m}\frac{du}{dV} \qquad \text{from (2) and (1)}$$

If we write $p = dV/du$ the equation for V takes the Clairaut form

$$V = up + A/p, \qquad A = Nh^3/8\pi^2 m \qquad (9)$$

The solution is

$$V = Cu + A/C, \qquad C = \text{integration constant} \qquad (10)$$

$$V = \frac{C}{r^2} + \frac{A}{C}, \qquad F = -\frac{dV}{dr} = \frac{2C}{r^3} \qquad (11)$$

It appears therefore that the force should be as the inverse cube and involves a (negative) constant C as yet undetermined.[3]

If we eliminate w between (1) and (2) we have

$$mr^3 = -n^2h^2/4\pi^2F = -n^2h^2r^3/8\pi^2C$$

Hence it appears that r cancels out and the quantum condition imposes

no restriction on the size of the orbit—a fact well known for the law of the inverse cube. Then

$$-C = n^2h^2/8\pi^2m = An^2/Nh \tag{12}$$

The force and potential energy therefore are

$$F = -\frac{2An^2}{Nhr^3} = -\frac{h^2n^2}{4\pi^2mr^3} \tag{13}$$

$$V = -\frac{An^2}{Nhr^2} - \frac{Nh}{n^2} = -\frac{h^2n^2}{8\pi^2mr^2} - \frac{Nh}{n^2} \tag{14}$$

It appears therefore that the force and potential energy depend on the quantum number; it is the force that is quantized and not the orbit.[4]

The differential equation (9) has, however, another solution, namely, the singular solution obtained by eliminating the constant between (10) and its derivative with respect to the constant.

$$V = \frac{C}{r^2} + \frac{A}{C} \qquad 0 = \frac{1}{r^2} - \frac{A}{C^2}$$

$$V = -\sqrt{\frac{Nh^3}{2\pi^2m}}\frac{1}{r} = -\frac{e^2}{r} \qquad \text{since } e^4 = \frac{Nh^3}{2\pi^2m}$$

This is Coulomb's law. From the general line of argument which leads to this result, it would appear that if we had given the quantum equations (2) and (3), the spectral law (4) and if we assume so much of elementary mechanics as the energy equation (5) and the force equation for circular orbits (1) and further assume that equation (8), therefrom deduced and known to hold for an infinity of orbits, holds identically, then we are led either to a quantized force (13) or to the Coulomb law.[5]

[1] Cf. G. N. Lewis, *Valence*, especially p. 42 and p. 50.

[2] To be sure, the experiments of Nicholson and Merton on the radii of the atomic orbits in mixtures of hydrogen and helium relative to the distance between molecules and of Ramsauer on the long free paths of electrons in argon and other gases do not favor Coulomb's law. It may be that the repulsion between nuclei and the attraction between nucleus and electron follow different laws.

[3] The inverse cube has been used in tentatives toward atomic theory by J. J. Thomson.

[4] We do not have here the quantized force of G. C. Evans, these PROCEEDINGS, 9, 1923, p. 234; for his term in the inverse cube represents repulsion and not attraction.

[5] This does not get us ahead much; but we may observe that of (1)–(5) only (4) is an experimental fact, (1) and (5) are definitions, (2) and (3) are hypotheses. There is some advantage in replacing the hypothetical Coulomb law by an experimental fact.

DENSITY AND DIFFUSION MEASUREMENT BY DISPLACEMENT INTERFEROMETRY IN EXTREME CASES[1]

BY CARL BARUS

DEPARTMENT OF PHYSICS, BROWN UNIVERSITY

Communicated June 24, 1924

1. *Behavior of Hydrogen.* The method described in the preceding paper[2] was applied for the case of a gas of small density (H_2); but the results came out much too large and the diffusion coefficients correspondingly too small. Nevertheless the data in successive repetitions of the diffusion experiments agreed very well with each other. In Figure 1, I give two examples among many, in which the ordinates s are the fringe displacements ($\Delta x = 21.5 \times 10^{-6} \Delta s$, Δx being the slide micrometer displacement corresponding to Δs) observed at the times in minutes given by the abscissas. One of the diffusions is prolonged for nearly an hour. The presence of H_2 in the open tube (68.4 cm. long) is still apparent after several hours.

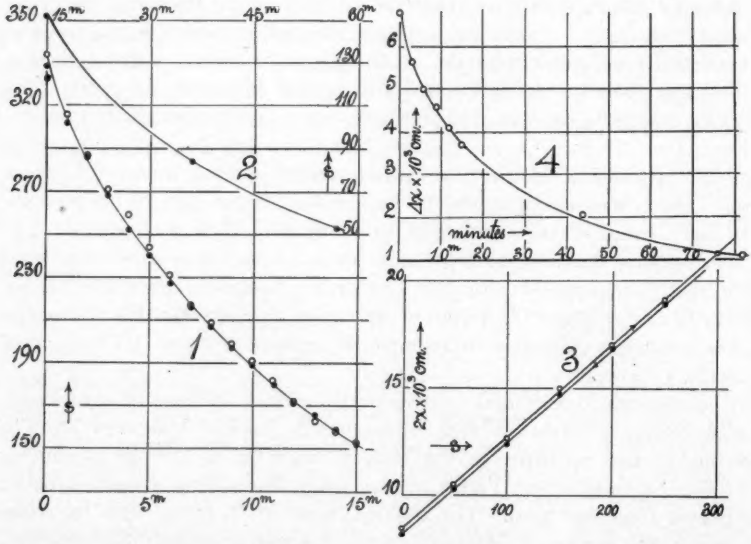
Careful consideration of the various sources of error involved in these gas density measurements, led to the conclusion that part of the pressure in the U-gauge is taken up by the surface tension of the mercury and does not appear in the mercury depression $\Delta h = \Delta x \cos \theta$, $\theta = 46.8°$, registered by the interferometer. In fact computing backward from the known density of hydrogen, the values of Δx are on the average 0.00077 cm. too small, which is equivalent to an error in pressure of $\Delta h = 0.00054$ cm. of mercury about.

Fearing that there might be some irregularity in the relation of measurements made in terms of fringe displacements (s ocular scale parts) and in terms of displacements of the slide micrometer of the interferometer (x cm.), a comparison, Figure 3, was made throughout a much broader interval than was used. The relation, however, is linear with the same slope in the ascending series and in the descending series. Nevertheless, there is a lag, unexplained, between the two. Consequently, subsequent measurements were made in terms of x only, the slide micrometer being previously set in place and the fringes used only to determine the zero or fiducial position.

2. *Fresh Mercury.* As the mercury of the U-gauge had become old and tarnished, these discrepancies were ascribed to it and a new charge of fresher distilled mercury was therefore introduced into the gauge, and other minor improvements added. Using the same vertical quill tube (length 68.4 cm., diameter 0.3 cm.), the mean of five experiments gave practically the same large density ($\rho_0 = 181 \times 10^{-6}$, as compared with $\rho_0 = 188 \times 10^{-6}$, above) as before. Thus the value of Δx was again

deficient by 7.1×10^{-4} cm., implying that a pressure of 5×10^{-4} cm. of mercury is borne by the capillary forces and not registered in depression of the surface of the U-gauge. The loss of pressure owing to the diffusion of hydrogen out of the tube, now measured by the slide micrometer only ($\Delta h = \Delta x \cos 45°$ cm. of mercury), is shown in Figure 4, throughout about 80 minutes.

If the discrepancy in question is in the gauge, then by using longer tubes so as to reduce the percentage effect of the error, the density of hydrogen obtained should approach the normal value. A variety of experiments made, fully bore this out. Thus in case of a tube of eighth inch gas pipe (diameter 1.6 cm.) 158 cm. long, the density obtained was $\rho_0 = 130 \times 10^{-6}$. Half-inch gas pipe, 70 cm. long, reproduced the former larger value.



It follows therefore that for pressures below 5×10^{-4} cm. of mercury, little can be done but to add to the micrometer displacement, Δx, the capillary error δx. The equations for the density at $t°$ should therefore read (ρ_t to be corrected for vapor pressure)

$$\rho_t = \rho_a - C(\Delta x + \delta x) - \rho_w$$

where ρ_w is the density of water vapor at $t°$, if present, ρ_a the density of air and $C = \rho_m \cos \theta / H$ for the tube length H and interferometer angle θ, ρ_m being the density of mercury. In the last example, if $\delta x = 6 \times 10^{-4}$ cm. be taken

$$\rho_t = 0.001192 - 0.0608(0.0173 + 0.0006) - 0.000017 = 83 \times 10^{-6}$$

whence $\rho_0 = 89.7$ happens to be exactly the density of hydrogen. This would fall off about 2.5% on correcting for vapor pressure.

Coefficient of Expansion of a Gas. The occurrence of a static corrective δx referred temporarily to the surface tension of mercury is surprising. It was therefore determined to test the matter further by measuring the density of air heated to 100°, or in other words, to find the coefficient of expansion. Here, for example, with an eighth inch gas pipe 70 cm. long, surrounded by a steam jacket while the outside air temperature was 27°, the slide micrometer showed $\Delta x = 0.0015$ cm. Using as above $\delta x = 0.0006$ for the static error $\Delta\rho = 0.137\ (\Delta x + \delta x) = 288 \times 10^{-6}$; $\Delta\rho/\Delta t = 3.95 \times 10^{-6}$ and $(\Delta\rho/\rho_0)/\Delta t = 0.0037$, a close approach to the coefficient of expansion, where Δx alone would have given a value of 40% too low.

[1] Advance note, from a Report to the Carnegie Institution of Washington, D. C.

[2] These PROCEEDINGS, **10**, 1924, p. 153–5.

IONIC MOBILITIES IN GASEOUS MIXTURES

BY L. B. LOEB AND M. F. ASHLEY

DEPARTMENT OF PHYSICS, UNIVERSITY OF CALIFORNIA

Communicated June 3, 1924

Since 1909 there have been two rival theories[1,5,6] of the nature of the normal gas ions purporting to describe their behavior. One theory assumes the ion to be a cluster of neutral molecules about a charged molecule, and constitutes the so-called "cluster" ion theory. On this theory the mobility is decreased below that to be expected for a charged molecule which does not exert forces on the surrounding molecules, because of the decreased mean free path of the large cluster ion. The alternative, or "small ion" theory assumes that the ion is but a single charged molecule whose average free path is decreased by its attractive forces on the surrounding neutral molecules. Many attempts have been made by means of experiments to decide between them with but little success. In 1916[2] one of the writers performed what he considered a crucial experiment which, while it gave results that appeared to favor the small ion theory was not the experimentum crucis anticipated. With the hope of getting further evidence on this question, the writers set themselves the task of determining ionic mobilities in mixtures of ammonia gas and air.

The mobility of the positive ion in NH_3 is given as about 0.6 cm./sec. per volt/cm. while that in air is 1.8 cm./sec. The assumption that this difference in mobility is due to a clustering effect would lead one to expect

the mobility of the ions in air to drop down to a low value for ammonia as soon as appreciable quantities of NH_3 were added, provided that the mobility measurements were made in such a manner that the cluster about the ion had time to form. On a small ion theory it might at first sight be assumed that if the decrease in mobility in NH_3 were due to some influence of the NH_3 molecules on the mean free path of the ions, the mobility would change from that of air to that of ammonia in a linear fashion with the quantity of NH_3 present.

To investigate this, mobility measurements in dry NH_3 and air mixtures were made for positive and negative ions. The Franck[3] modification of the Rutherford alternating current method utilizing a square wave form alternating potential from a commutator was employed. The ionization chamber and commutator[3] have been described elsewhere. The procedure was to exhaust the ionization chamber to a pressure of less than .001 mm., then to fractionate into it to the required partial pressure NH_3 gas which came from a tube containing solid or liquid NH_3. Dry air was then admitted to bring the total pressure up to atmospheric pressure. This air was passed through a liquid air trap before it entered the chamber, the air having first been dried by passing through a drying train previously described.

Two sources of NH_3 were used. The first was a small tank of commercial liquid NH_3, the later supply came from a concentrated solution of NaOH acting on chemically pure NH_4Cl. The gas was dried by passing through a tube 150 cms. long filled with small granules of NaOH. Before each filling this tube was dried by prolonged pumping with a mercury diffusion pump to a pressure of less than .01 mm. After passing over these tubes the NH_3 was condensed out by liquid air. The supply tank was shut off and any traces of air that came over with the gas were pumped out while the NH_3 was frozen at liquid air temperature. The liquid air was next removed and the gas distilled into the ionization chamber through two traps cooled to —18°C. to remove any traces of water vapor. Care was taken to use only the first three fourths of the liquid NH_3 in the tube for a filling, since under these conditions the temperature of the liquid NH_3 was quite low and very little moisture was carried over into the ionization chamber. Out of 30 cc. of the liquid NH_3 about $^1/_{10}$ of a cc. of water was left in the bulb.

The plate distance in the measurements was 1.192 cm. and an alternating potential of about 30 cycles per second was used throughout. The value of the mobility, however, was not changed by changing the frequency, the frequency being kept constant to reduce all possible variables.

The results obtained for positive and negative ions plotted as ordinates against the per cent of NH_3 present are shown as the circled points and crosses in Fig. 1.* The curves (1) represent the expected linear change for the

small ion theory, while a curve of the type of (2) represents what might be expected on an extreme cluster theory point of view. It is seen that again the results are indecisive.

In a paper published in 1917 the writer showed that the mobility of the gaseous ions in different gases followed a modified form of Kaufmann's rule.[1] This rule states that the mobility of an ion in one gas varies inversely as the square root of the molecular weight of the gas, and in this form it is applicable to the more "perfect" gases only. The writer showed that it was more or less applicable to all gases and vapors if the dielectric

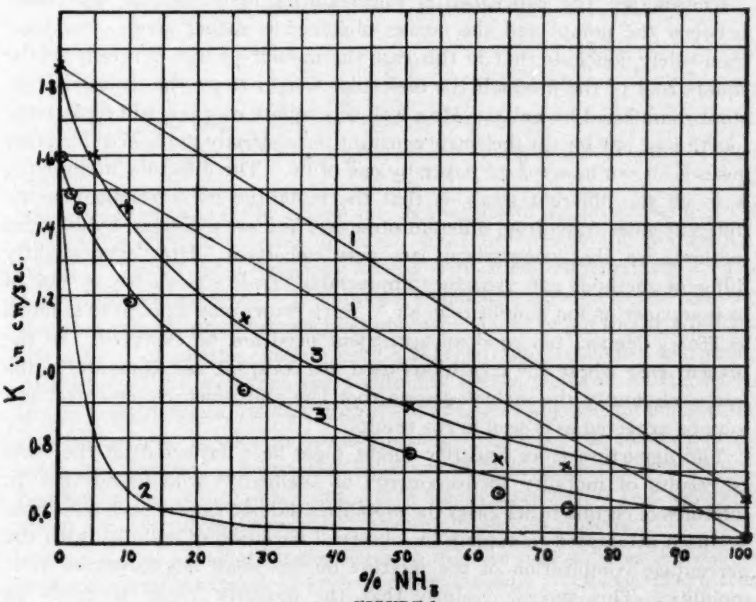


FIGURE 1

constant of the vapor or gas were also taken into account. This means that all mobilities of ions in gases could be reduced to the value of the mobility of the gases in hydrogen by multiplying them by the square root of their molecular weight relative to hydrogen and by multiplying them by the fourth root of the dielectric constant of the gas diminished by unity relative to that for hydrogen. This at the time was purely an empirical rule having no apparent justification in theory. An attempt was made to apply this type of rule to the mobilities of ions in the air-ammonia mixtures studied, with the following result: If one plot the value of Kc as computed from the equations below against c the per cent of ammonia gas present, one obtains the full curves 3 and 3, Fig. 1.

$$K_{NH_3} = K_{air} \frac{1}{\sqrt{\frac{M_n(D_n - 1)}{M_a(D_a - 1)}}} = \frac{K_{air}}{\sqrt{Q}}$$

where D_n and D_a are the dielectric constants of NH_3 and air, and M_n and M are the molecular weights of NH_3 and air.

$$K_c = \frac{K_{air}}{\sqrt{\frac{(100 - c)1 + Qc}{100}}}$$

Considering the experimental uncertainties involved, the agreement between the points and the curves observed is rather good. One may then safely conclude that in this case the mobility varies inversely as the square root of the product, the molecular weight times the dielectric constant diminished by unity. How well this square root law will replace the fourth root law for the dielectric constant term already published[1] for other gases is shown in a second paper by one of us. The difficulty in applying it to all the different gases, is that the mobilities as determined in the different gases come from different observers and are measured by different methods, so that comparisons are very uncertain. How even slightly different methods can introduce appreciable errors was shown by one of us in a paper on ion mobilities in air.[3] Such errors may amount to as much as 20%. Hence, too good an agreement need not be expected. In the present case where the conditions used are constant and comparable the law is obviously the inverse square root law and much closer agreement may be expected as is seen in the results.

The departure from linearity might have been expected on the basis of results of mobility measurements of Wellisch[4,5] who found that in mixtures of certain other gases the mobilities did not vary in such a fashion. Wellisch attempted to explain his observed variation of mobility with the percentage composition of the mixture on his small ion theory of ionic mobility. This theory assumes that the mobility varies inversely as $(1 + b)(D-1)$, that is, approximately, inversely as the first power of $D-1$. This is not in agreement with the present experiments nor with some of Wellisch's[5] experiments on mobilities in gaseous mixtures.

In a recent paper, J. J. Thomson[7] has worked out a theory of the decrease of the mean free path of the ion in moving through a gas on more rigorous grounds than has been done before using an inverse fifth power law of force. His change of mean free path due to the charge, which is $\lambda' = \frac{\lambda}{2.2\sqrt{\frac{2k}{\sigma^4 Mc^2}}}$ † may be inserted into the mobility equation of Langevin[8,9] for an ion as may be seen below. This causes the mobility to vary in-

versely as the square root of the product of the mass of the molecules of the gas in which the ion moves into their dielectric constant diminished by unit. Here k is the force constant, which for a gas of dielectric constant D could be written $\frac{(D-1)e^2}{2\pi n_0}$, where n_0 is the number of molecules per cc. and e is the ionic change. Langevin's equation for the mobility may be written:

$$K = .815 \frac{e}{M} \frac{\lambda'}{C} \sqrt{\frac{M+m}{m}}$$

Inserting Thomson's λ' and setting $m = bM$, where b is the number of molecules in the cluster one obtains:

$$K = .815 \frac{e}{M} \frac{\lambda}{C} \frac{\sqrt{\frac{b+1}{b}}}{2.2\sqrt{\frac{(D-1)e^2}{\pi\sigma^4 n_0 Mc^2}}} = \frac{.371\,\lambda\sqrt{\frac{b+1}{b}}}{\sqrt{\frac{(D-1)M}{\pi n_0 \sigma^4}}};$$

where m is the mass of the cluster ion, M the mass of the gas molecule (σ the sum of the radii of molecule and ion), and c the velocity of thermal agitation. Thus it is in excellent accord with these measurements. The particular treatment of J. J. Thomson's equation here given and its consequences which are of far reaching importance for the ion theory will form the subject of another paper being published simultaneously elsewhere. In this it will be seen that the Kaufmann law as modified by the writer in the present paper, follows from the Thomson-Langevin mobility equation and is in no way a criterion distinguishing between the cluster and the small ion theories. In fact this equation as modified is practically independent of the nature of the ion. It is based, however, as J. J. Thomson shows, on a set of conditions which eventually lead to clustering.

* The absolute values of the mobilities given here are those observed corrected for what was essentially a field of +2.5 volts applied to the gauze. This seems to have come from a contact difference of potential resulting from the action of the NH_3 vapor on the gauze and the plates which were of dissimilar metals.

† The Mc^2 here used is related to Thomson's quantity $\frac{mMV^2}{M+m}$ by the relation $\frac{mMV^2}{m+M} = Mc^2$. Where V in Thomson's equation is the relative velocity of ion and molecule.

[1] Loeb, L. B.: *J. Franklin Inst. Phila.*, **184**, 775, Dec. 1917.
[2] Loeb, L. B.: *Physic. Rev.*, **8**, 633, Dec. 1916.
[3] Loeb, L. B.: *Ibid.*, **17**, 69, Feb. 1921, and *J. Franklin Inst.*, **196**, 537, Oct. 1923.
[4] Wellisch, E. M.: *Proc. Roy. Soc.*, **A82**, 500, 1909.
[5] Wellisch, E. M.: *Phil. Trans. Roy. Soc.*, **A209**, 249, 1909.
[6] Sutherland, W.: *Phil. Mag.*, **18**, 341, 1909.
[7] Thomson, J. J.: *Ibid.*, **47**, 342, Feb. 1924.
[8] Langevin, P.: *Ann. Chim. Phys.*, (5) **8**, 245, 1905.
[9] Mayer, H. F.: *Ann. Physik.*, No. **12**, 1920.

CONCERNING UPPER SEMI-CONTINUOUS COLLECTIONS OF CONTINUA WHICH DO NOT SEPARATE A GIVEN CONTINUUM

BY R. L. MOORE

DEPARTMENT OF PURE MATHEMATICS, UNIVERSITY OF TEXAS

Communicated, June 30, 1924

In her thesis,[1] Miss Anna M. Mullikin showed that if M is the set of all points in a given plane and G is a countable collection of mutually exclusive closed subsets of M no one of which disconnects[2] M then M is not disconnected by the sum of all the point sets of the collection G. It is clear that this proposition does not remain true if the phrase "M is the set of all points in a given plane" is replaced by the phrase "M is a closed and connected subset of the set S of all the points of a given plane." If M is any simple closed curve, A and B are two points belonging to M and the point sets of the collection G are the points A and B then no point set of the collection G separates[2] M, but their sum does separate M. It is to be observed that in this case M disconnects S. Consider, however, the following example.

Example.[3] For each n, let d_n denote the arc composed of the four straight line intervals A_nB_n, B_nC_n, C_nD_n and D_nE_n, where A_n, B_n, C_n and D_n are the points $(1/2^{n-1}, 0)$, $(1/2^{n-1}, 1/2^{n-1})$, $(-1/2^{n-1}, 1/2^{n-1})$, $(-1/2^{n-1}, -1/2^{n-1})$ and $(1, -1/2^{n-1})$ respectively. Let d denote the straight line interval from (0,0) to (1,0). Let G denote the collection of point sets $d_1, d_2, d_3, \ldots$ and let M denote the point set obtained by adding together d and all the point sets of the collection G. Clearly M is closed and bounded, no point set of the collection G separates M and M does not separate S. But, nevertheless, the sum of all the point sets of the collection G does separate M. It is to be observed, however, that G is not an upper semi-continuous collection in the sense indicated below.

Definition. A collection of continua G is said to be an *upper semi-continuous* collection provided it is true that if p is a continuum of the collection G and $p_1, p_2, p_3, \ldots$ is a sequence of continua of the collection G such that the lower[4] distance from p_n to p approaches zero as a limit as n approaches infinity then so does the upper[4] distance from p_n to p.

THEOREM 1. *If, in a plane S, M is a bounded continuum which does not separate S and G is a countable upper semi-continuous collection of subcontinua of M no one of which separates M, then the sum of all the continua of the set G does not separate M.*

In the proof of Theorem 1 given below use will be made of three lemmas.

LEMMA 1. *If the points A and B of the bounded continuum M are sepa-*

rated in M by a point set K then they are separated in M by some closed subset of K.

Proof. By hypothesis, $M-K$ is the sum of two mutually separated point sets M_A and M_B which contain A and B respectively. By a theorem of Knaster and Kuratowski's,[5] there exists a continuum H which contains no point either of M_A or of M_B but which separates A from B in S. Let L denote the set of points common to M and H. Since M and H are closed, L is closed. Clearly L separates A from B in M.

LEMMA 2. *If G is an upper semi-continuous collection of continua and K is a closed set of points and each point of the set K belongs to some continuum of the collection G then the sum of all those continua of the collection G which have points in common with K is a closed set of points.*

The truth of Lemma 2 may be easily established.

LEMMA 3. *If H and K are two bounded continua whose common part T is not connected then there exists a bounded domain which is complementary to $H + K$ and whose boundary contains at least one point of T and at least one point of $H-T$.*

Proof. If there exists a bounded domain which is complementary to $H + K$ and whose boundary is a subset of K then let K^* denote the point set obtained by adding every such domain to K. Otherwise let K^* denote K. If there exists a bounded domain which is complementary to $H + K$ and whose boundary is a subset of $H-T$ let H^* denote the point set obtained by adding every such domain to H. Otherwise let H^* denote H. Clearly H^* and K^* are bounded continua and T is their common part. But, by hypothesis, T is not connected. It follows, by a theorem of Janiszewski's,[6] that $H^* + K^*$ separates S and therefore there exists at least one bounded domain R which is complementary to $H^* + K^*$. Evidently R is complementary to $H + K$. But the boundary of R is not a subset of $H-T$ or of K. For if it were then R would be a subset H^* or of K^* which is impossible since it is complementary to $H^* + K^*$. Hence the boundary of R contains at least one point of $H-T$ and at least one point of K and therefore, since it is[7] connected, it contains a point of T.

Proof of Theorem 1. Suppose, on the contrary, that M is disconnected by the point set T obtained by adding together all the continua of the set G. By Lemmas 1 and 2 the collection G contains a subcollection G^* such that T^*, the sum of all the point sets of the collection G^*, is a closed point set which disconnects M. The point set $M-T^*$ is the sum of two mutually separated point sets H and K. For each point X in H let H_X denote the greatest connected subset of $H + T^*$ which contains X. It is clear that, for every X in H, H_X contains at least one point set of the collection G^* and if it contains one point of a point set of this collection then it contains the whole of that point set. Clearly there exists, in H, a countable set of points $X_1, X_2, X_3, \ldots$ such that every point of H be-

longs to some point set of the collection $H_{X_1}, H_{X_2}, H_{X_3}, \ldots$ and such that no two point sets of this collection have any point in common. There are two cases to be considered.

Case 1. Suppose that, for every n, H_{X_n} contains a limit point of $(H + T) - (H_{X_n} + T)$. Let U denote the collection of point sets H_{X_1}, H_{X_2}, H_{X_3}, ..., together with all those continua of the collection G^* which do not belong to any H_X but which contain one or more limit points of H. The sum of all the continua of the countable collection U is closed and bounded and each continuum of this collection contains at least one limit point of the sum of all the others. But, according to a theorem which I have recently[8] established, there exists no countable collection of continua satisfying these conditions. Thus, in Case 1, we have a contradiction.

Case 2. Suppose that there exists a point set E which belongs to the collection $H_{X_1}, H_{X_2}, H_{X_3}, \ldots$ and which contains no limit point of $(H + T) - (E + T)$. For every point Y of the set $M - E$ let D_Y denote the greatest connected subset of $(M - E) + T^*$ which contains Y. It is clear that, for every Y, D_Y is closed and that D_Y and E have, in common, at least one continuum of the collection G^*. They cannot have, in common, more than one continuum of this collection. For suppose they do. Then their common part (call it N) is the sum of a countable number of (two or more) mutually exclusive continua and therefore, since it is closed, it is not[9] connected and consequently $D_Y + E$ separates the plane, and, by Lemma 3, there exists a bounded domain R which is complementary to $D_Y + E$ and whose boundary β contains at least one point of N and at least one point of $E - N$. The point set β is not a subset of T^*. For if it were then, since it contains at least one point of N and at least one point of $E - N$ and N is a subset of T^* and every point set of G^* which contains a point of N belongs wholly to N therefore β would be the sum of a countable number of (two or more) continua of the set G^*, which is impossible,[8] since β is a bounded continuum. It follows that β contains at least one point P^* of E which does not belong to T^* and therefore belongs to H. There exists a circle C with center at P^* which encloses no point of $K + T^*$. Every point of β within C belongs to H. But there exists within C a point P which belongs to β and which is accessible from R, that is to say which is an end-point of some arc PX which lies, except for the point P, wholly in R. Since M does not separate S, every point in R, and therefore every point on the arc PX, belongs to M. Since P belongs to H and H contains no limit point of $K + T^*$ therefore there exists, on the arc PX, a point Z such that the interval PZ of the arc PX contains no point of $K + T^*$ and therefore is a subset of H. But this is impossible, since R contains no point of E and E is a maximal connected subset of $H + T^*$. Thus the supposition that D_Y and E have, in common,

more than one continuum of the collection G^* has led to a contradiction. Hence the set of points common to D_Y and E is a single continuum of the collection G^*. Also, for every two points Z and W of the set $M-E$, either D_Z and D_W are identical or they have no point in common. It follows that there exist, in $M-E$, a countable number of points $Y_1, Y_2, Y_3, \ldots$ such that every point of $M-E$ belongs to some point set of the collection $D_{Y_1}, D_{Y_2}, D_{Y_3}, \ldots$ Let V denote the collection of point sets consisting of $D_{Y_1}, D_{Y_2}, D_{Y_3}, \ldots$ together with all those continua of the collection G^* which are subsets of E and do not belong to any D_Y, but which contain one or more limit points of $M-E$. It is clear that every point set of the collection V which is not a D_Y contains at least one limit point of the sum of all the other point sets of the collection V. Suppose that, for some Y, D_Y does not contain a limit point of the sum of all the other point sets of the collection V. Let g_Y denote that continuum of the collection G^* which is common to D_Y and E. It is clear that no point of D_Y-g_Y is a limit point of $M-D_Y$. Furthermore, no point of $M-D_Y$ is a limit point of D_Y-g_Y. But $M-g_Y = (M-D_Y) + (D_Y-g_Y)$. Therefore $M-g_Y$ is not connected. But this is contrary to hypothesis. Hence every point set of the countable collection V contains a limit point of the sum of all the other point sets of that collection. Furthermore, the sum of all the point sets of the collection V is closed and bounded. Thus we have a result which is in contradiction to a theorem[8] mentioned above under Case 1. The truth of Theorem 1 is therefore established.

THEOREM 2. *If, in a plane S, M is a bounded continuous curve which does not separate S, and A and B are two distinct points of M, and G is a countable upper semi-continuous collection of subcontinua of M no one of which separates A from B in M, then the sum of all the continua of the set G does not separate A from B in M.*

Proof. Suppose this is not the case. Then the collection G contains a subcollection G^* such that T^*, the sum of all the continua of the collection G^*, is a closed point set which separates A from B in M. The point set $M-T^*$ is the sum of two mutually separated point sets H and K which contain A and B respectively. Let H_A denote the greatest connected subset of $H + T^*$ which contains A. There does not exist, in H_A, a point P not belonging to T^* such that P is a limit point of $M-H_A$. For if such a point existed M would clearly not be connected im kleinen at that point and therefore M would not[10] be a continuous curve. It may be shown, in a similar way, that if K_B denotes the greatest connected subset of $(M-H_A) + T^*$ which contains B then K_B contains no limit point of $M-K_B$ which does not belong to T^*. By a method similar to that used in the treatment of Case 2 under the proof of Theorem 1, it may be shown that the common part of H_A and K_B is some single continuum g of the set G^*. It is clear that the point sets K_B-g and $M-(K_B-g)$

are mutually separated. But these point sets contain A and B respectively and their sum is the point set $M-g$. Therefore g separates A from B in M. But this is contrary to hypothesis. The truth of Theorem 2 is therefore established.

If, in the statement of Theorem 2, the phrase "continuous curve" is replaced by the word "continuum," the resulting proposition is false. That this is the case may be seen with the help of the following example.

Example. Let A denote the point (0,0), let B denote the point (1,0) and, for each positive integer n, let B_n denote the point $(1,1/n)$. Let M denote the continuum obtained by adding together the points of the intervals AB, AB_1, AB_2, AB_3, ... Let C denote the point (1/2,0) and, for each n, let C_n denote the point $(1/2,1/2n)$. Let G denote the countable collection of point sets C, C_1, C_2, C_3, ... In this case each point set of the collection G consists of a single point and therefore G is necessarily an upper semi-continuous collection. Clearly, M does not separate S and no point set of the collection G separates A from B in M. But A is separated from B in M by the sum of all the point sets of the collection G.

[1] "Certain Theorems Relating to Plane Connected Point Sets," *Trans. Amer. Math. Soc.*, **24**, 144–62(1922).

[2] Two point sets are said to be *mutually separated* if they have no point in common and neither of them contains a limit point of the other one. A point set is said to be *connected* if it is not the sum of two mutually separated point sets (Cf. N. J. Lennes, *Amer. Journ. Math.*, **33**, 1911). The point set H is said to *disconnect* the connected point set M if H is a proper subset of M and $M-H$ is not connected. If H, K and L are three mutually exclusive subsets of a connected point set M, then K is said to *separate H from L in M* provided $M-K$ is the sum of two mutually separated point sets which contain H and L respectively.

[3] This is a modification of an example constructed, for another purpose, by Miss Mullikin, *loc. cit.*, p. 145.

[4] If A and B are two points let $r(AB)$ denote the distance from A to B. If A is a point and q is a continuum, the *lower distance from A to q*, denoted by $l(Aq)$, is the lower bound of the distances $r(AY)$ for all points Y of q. If p and q are continua the *lower distance from p to q* is the lower bound (and the *upper distance from p to q* is the upper bound) of the distances $l(Xq)$ for all points X *of* p. It is to be observed that the upper distance from p to q is not necessarily the same as the upper distance from q to p.

[5] B. Knaster et C. Kuratowski, "Sur les ensembles connexes," *Fundamenta Mathematicae*, **2**, 233(1921).

[6] S. Janiszewski, "Sur les coupures du plan faites par des continus," *Prace matemfizyczne*, **26**, 1913.

[7] L. E. J. Brouwer, *Math. ann., Leipzig*, **69**.

[8] R. L. Moore, "An extension of the theorem that no perfect point set is countable," these PROCEEDINGS, **10**, 168–170(1924).

[9] Cf. W. Sierpinski, "Un theoreme sur les continus," *Tohoku Math. Journ.*, **13**.

[10] Cf. Hans Hahn, *Wien. Ber.*, **123**, 2433(1914). Also S. Mazurkiewicz, *Fundamenta Mathematicae*, **1**, 166–209(1920).

THE SPECTROHELIOSCOPE

BY GEORGE E. HALE

MOUNT WILSON OBSERVATORY, CARNEGIE INSTITUTION OF WASHINGTON

Communicated, July 7, 1924

Monochromatic images of the sun, photographically recorded with the spectroheliograph, reveal the phenomena of the solar atmosphere in projection against the disk. The light from a spectral line of calcium, hydrogen, or other substance is singled out by a narrow second slit, which moves across the plate while the first slit moves across the solar image. A monochromatic picture is thus built up of countless narrow slit images, recorded side by side in slow succession.

The vortices and other significant structures thus disclosed, with the exception of an occasional brilliant eruption or unusually dark hydrogen flocculus, are beyond the reach of visual observation with the spectroscope. The simple method of opening the slit, which affords an excellent view of prominences at the sun's limb, because the light of the sky is sufficiently weakened by dispersion, is seriously limited when applied to the intensely brilliant disk. Even with the highest dispersion the slit cannot be opened sufficiently to reveal the characteristic structure of the dark hydrogen flocculi. Moreover, with a widened slit the image is not strictly monochromatic, and differences in form shown by the spectroheliograph with light from different parts of Hα obviously could not be detected in this way.

The simple expedient of rapidly oscillating the narrow spectroscope slit, and synchronously with it a second slit transmitting the Hα line to the eye, was tried by the earliest observers of the prominences. With suitable precautions this device gives excellent images of prominences through persistence of vision, but it did not survive the introduction by Zoellner and Huggins of the wide slit method, and apparently was not tried for the observation of phenomena on the sun's disk.

Many years ago I made some preliminary trials of the oscillating slit method with the 30-foot spectroscope and 60-foot tower telescope on Mount Wilson. I have only recently found opportunity to continue these experiments and to develop a satisfactory instrument based on this principle with which I have now secured good visual observations of both bright and dark flocculi.

An image of the sun two inches in diameter, given by a coelostat and a 12-inch objective formerly belonging to the Kenwood Observatory,* was observed with a grating spectroscope mounted horizontally. In this instrument the light passing through the slit falls on a 6-inch concave mirror of 13-feet focal length, which returns a parallel beam to a point

just below the slit, where a 6-inch plane grating is mounted. This sends the diffracted beam to a second 6-inch concave mirror, supported below the collimating mirror, which forms an image of the spectrum in the same plane as the first slit, and immediately below it. A fixed second slit at this point permits any part of the spectrum to be isolated. The grating was ruled by Jacomini with about 15,000 lines to the inch on the ruling machine of the Mount Wilson Observatory, with a diamond ground after Anderson's formula so as to give great brightness at the red end of the first order spectrum. The definition is perfect, and the brightness near Hα remarkable, as the attempt to concentrate most of the light in one spectrum and to favor the red end proved successful.

Suppose the first and second slits to be carried at opposite ends of a brass bar, mounted on a bearing half way between them and thus free to revolve about this center. Place the bar vertical, and turn the grating until the Hα line in the bright first order is bisected by the second slit. With the optical arrangement employed, a small displacement of the first slit to the right causes an exactly equal displacement of the Hα line to the left. Thus if the bar is oscillated back and forth by means of a small electric motor, a monochromatic image of the sun will be seen through a low power positive eyepiece focussed on the second slit.

This arrangement serves very well for the observation of prominences at the limb, where they can be seen at their full height with slits of moderate width. It also shows exceptionally bright or dark flocculi on the disk, though the slits must be narrower in order to give sufficient purity and reduce the brightness of the continuous spectrum. For flocculi of ordinary intensity the best results have been obtained with the aid of multiple slits, five at each end of the bar, 0.003 inch wide and 0.08 inch apart. A fixed slit, slightly less than 0.08 inch wide, must be used behind the upper slits, to prevent the formation of overlapping spectra. Two complete oscillations of the bar per second, corresponding to twenty illuminations of the retina, give a sufficiently steady image. A rotating disk, carrying a large number of radial slits, is in some respects a more satisfactory device for the same purpose.

This instrument, which may appropriately be called a spectrohelioscope, should prove a valuable auxiliary of the spectroheliograph in several kinds of work. It will permit the rapidly changing forms of eruptions on the disk to be followed visually, and be of special service in deciphering the curious differences of structure sometimes found on photographs of flocculi taken simultaneously with the opposite edges of Hα. As the oscillating bar can be moved toward red or violet by a micrometer screw while observations are being made, the possibility of passing instantly from one edge of the line to the other should assist in the interpretation of the spectroheliograph results.

Further details, with a photograph of the instrument, will be published in the *Astrophysical Journal.*

* Kindly loaned me by Professor Frost. The apparatus was set up temporarily in my garden at South Pasadena.

THE CONSTITUTION OF THE SO-CALLED "PECHMANN DYES" AND THE MECHANISM OF THEIR FORMATION FROM BETA-BENZOYLACRYLIC ACID[1,2]

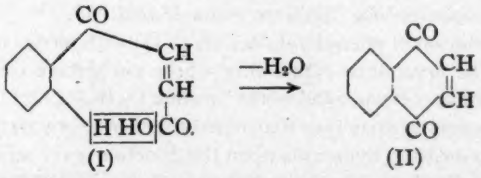
BY MARSTON TAYLOR BOGERT AND JOHN J. RITTER

ORGANIC LABORATORIES, COLUMBIA UNIVERSITY

Communicated, July 21, 1924

For 42 years the constitution of the so-called "Pechmann Dyes" has remained an unsolved problem. Based upon experimental facts observed by us, the conclusions presented in the following pages are submitted as a contribution to the solution of this problem.

Von Pechmann,[3] who was the first to attempt the production of alpha-naphthaquinones (II) by dehydration of *beta*-benzoylacrylic acid (I) or its homologs,

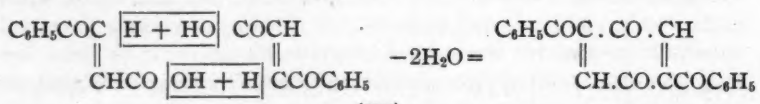


(I) (II)

obtained instead highly colored products which, altho of the same percentage composition as the expected naphthaquinones, were totally different and whose nature was not ascertained. These have since been known generally as "Pechmann Dyes." They are characterized by brilliant colors and striking color changes, as well as by other interesting physical and chemical properties. Their absorption spectra and certain of their color reactions resemble those of pigments related to chlorophyll; and by the study of their behavior with ammonia and with amines, we hope to throw some additional light upon the question of the structure of the coloring matters of the hemopyrrol and phyllopyrrol group, possibly involving also their immediate progenitors, such as hemin, hematoporphyrin, bilirubin, and others.

Kozniewski and Marchlewski,[4] 24 years later, repeated von Pechmann's work and extended it. Incidentally, they mentioned in a footnote the significant fact that the dye from mesitoylacrylic acid was obtained always

in two different forms, one red and the other yellow. They came to the conclusion that these dyes were diaroylquinones, formed by a dehydration involving the beta hydrogen atom,



(III)

basing this assumption upon the deep color of the compounds, the ease with which they yielded di-anils with aniline, and benzoic acid on oxidation.

We have found, however, that the *alpha*-methyl derivative of *beta*-benzoylacrylic acid does not give a Pechmann dye. Nor does the quinone formula provide any simple answer tothe question as to how it yields diphenacylfumaric acid by the action of alkali. It offers no explanation for the *cis-trans* isomerism observed by us, nor for the production of the dibromide obtained by Kozniewski and Marchlewski.

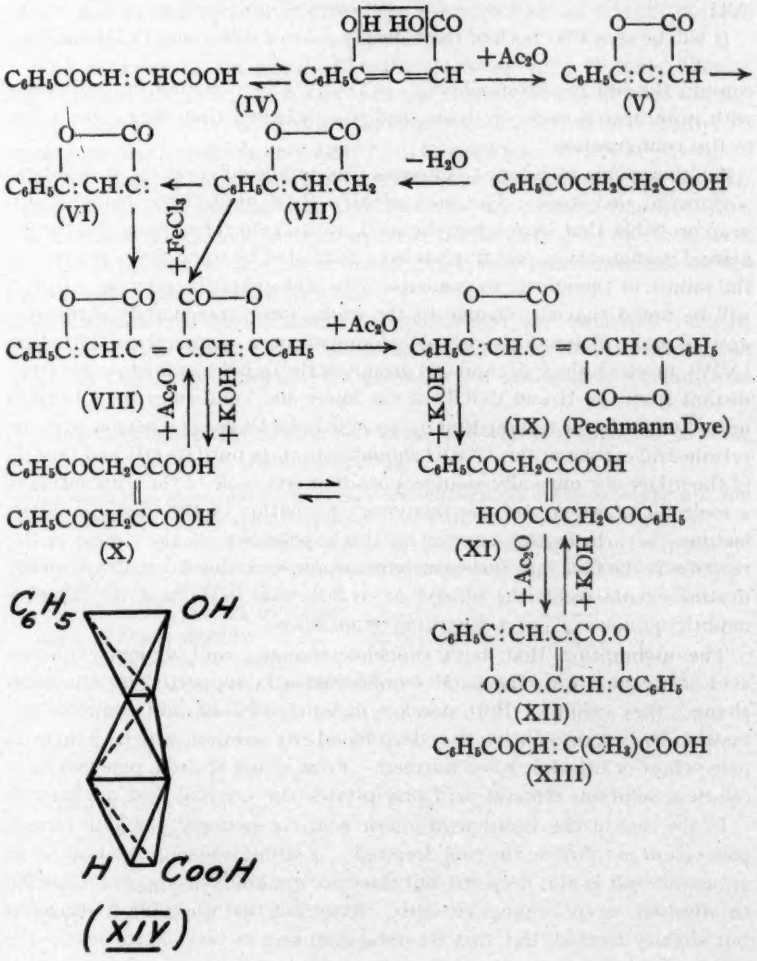
Von Pechmann stated that the action of alcoholic potassium hydroxide upon the dye yielded a yellowish product, which regenerated a red substance when boiled with acetic anhydride. Investigation of this yellow product has now shown it to be a mixture of a bright yellow acid, diphenacylfumaric (XI), with a colorless one, diphenacylmaleic acid (X).

By the oxidation of phenylcrotolactone (VII) with ferric chloride, Kugel[5] obtained a red crystalline compound, whose percentage composition and molecular weight corresponded to the formula $C_{20}H_{12}O_4$, and to which he assigned as a tentative structure that represented by Formula VIII. Action of alcoholic potassium hydroxide upon the dilactone gave an acid, $C_{20}H_{16}O_6$, which he suggested was probably diphenacylfumaric. When this acid was heated with glacial acetic acid, there was formed an isomer of the original dilactone, which he thought was probably a stereoisomer, and which he described as crystallizing in fine yellow prisms which appeared bright red in superimposed masses.

Kugel's dilactone is the diphenacylmaleic dilactone (VIII), and results from the simultaneous oxidation and union of two molecules of the phenylcrotolactone (VII). What he suggested might be diphenacylfumaric acid is actually that acid, and is identical with the yellow acid obtained by us from the Pechmann Dye and alcoholic potassium hydroxide. Its origin from a *cis*-dilactone is due to the ease with which this dilactone rearranges to the *trans* isomer. Dehydration of this acid, instead of yielding a dilactone (IX), gives the dilactide (XII), i.e., it closes to the 6- instead of the 5-membered cycle. On the other hand, the diphenacylmaleic acid (X), which should not form a dilactide, does not do so, but reverts directly to the Kugel dilactone (VIII) when it loses water.

This same *cis*-dilactone can be obtained by heating *beta*-benzoylacrylic acid for a short time with acetic anhydride. On longer heating of the benzoylacrylic acid with acetic anhydride, or on heating the *cis*-dilactone with the same reagent, it rearranges to the *trans* isomer, which is the Pechmann Dye. Depending upon the duration of the heating with acetic anhydride, therefore, the product will contain varying amounts of the two stereoisomers, hydrolysis of which by alcoholic potassium hydroxide will give corresponding quantities of diphenacylmaleic and diphenacylfumaric acids.

These different steps are exhibited in the accompanying flow-sheet:



Upon the foundation erected by the facts presented in the foregoing, we have constructed the following hypothesis to explain the mechanism of the formation of the Pechmann Dye from *beta*-benzoylacrylic acid: The first step consists in the transition of the acid to the enolic form (IV) with subsequent loss of water and transient production of the highly unsaturated lactone (V) of allene type, which promptly undergoes a bimolecular condensation to the Kugel dilactone (VIII), possibly through an unstable intermediate stage (VI) identical with what might be expected also in the production of the dilactone by Kugel's method of oxidizing the phenylcrotolactone (VII).

It will be seen that both of these deeply colored dilactones (VIII and IX) contain the indigo chromophore, —CO.C:C.CO—, united on both sides with unsaturated cyclic systems, and it is believed that their color is due to this configuration.

Another point of interest concerns the stereoconfiguration of *beta*-benzoylacrylic acid itself. The work of Mrs. Rice[6] and others, makes it appear probable that benzoylacrylic acid, in its keto form, possesses the fumaroid configuration, and this has been suggested by some as the reason for the failure of the efforts to condense it to *alpha*-naphthaquinone. But, it will be noted that the change to the enolic form immediately eliminates any *cis-trans*-distinction with development of a symmetrical structure (XIV), in which the C_6H_5 and OH groups of the upper tetrahedron are equidistant from the H and COOH of the lower one; and since the COOH is united to the lower tetrahedron by a single bond there is no reason why the tetrahedral carbon of this COOH should not rotate until its OH and the OH of the upper one mutually assume a position favorable to the elimination of a molecule of water and the temporary formation of the unstable allene lactone (V). In further support of this explanation of the course of the reaction is the fact that *beta*-benzoylpropionic acid when treated with dehydrating agents yields the phenyl-Δ^2-crotolactone (VII) and no dihydronaphthaquinone $\rightleftharpoons$ 1,4-dihydroxynaphthalene.

The assumption that both diphenacylfumaric and diphenacylmaleic acid easily pass into the enolic configuration is supported by the color changes they exhibit. Both dissolve instantly in cold dilute aqueous potassium hydroxide solution to a deep blood-red solution, which changes to pale yellow or colorless when warmed. From either the red, pale yellow or colorless solution, mineral acid precipitates the original acid unchanged.

In the case of the diphenacylfumaric acid, for example, the keto form is pale yellow in solution, the enol deep red. A strong aqueous solution of its potassium salt is also deep red, but this color gradually disappears following the dilution, as enol changes to keto. Assuming that the enolic hydrogen is but slightly ionized, but that its potassium salt in very dilute solution is

very largely ionized, and that the hydrolysis equilibrium adjusts itself practically instantly, the lag in decoloration may be taken as a measure of the time necessary for the transformation of the enolic to the ketonic form. The temporary red color of dilute aqueous solutions of the potassium salt, therefore, is probably to be ascribed to the free enolic acid. It should be possible to follow this change by measuring the decreasing conductivity of the salt solution, although a solution containing both carboxylic and enolic salts presents a rather complex system in which the degree of hydrolysis of either salt group will be affected by the extent of the hydrolysis of the other, due to alteration of hydroxyl-ion concentration as a result of such hydrolysis.

The behavior of the keto acid with alcoholic ferric chloride likewise discloses an interesting time factor in the rearrangement of keto to enol, for the characteristic red-brown color appears very tardily and reaches its maximum depth of shade only after 15 or 20 minutes.

Wolff[7] observed that *beta*-acetylacrylic acid dissolved in cold aqueous barium hydroxide solution with a yellow color, which color vanished when the solution was warmed, and interpreted this as indicating the formation of the enolic salt at the lower temperature and rearrangement to the colorless keto form at the higher one.

New members of the Pechmann Dye group have been prepared from the naphthoylacrylic acids. They are deeply colored crystalline solids, that from the *alpha*-naphthoylacrylic possessing a remarkable brilliant blue luster.

Other articles in this field are in press and the investigation is being continued actively in various directions.

[1] On account of space limitations on these Proceedings, the complete article, with full experimental details, will appear elsewhere—M. T. B.

[2] We welcome this opportunity to express publicly our appreciation of the courtesy of the Barrett Company who generously provided us with the maleic acid necessary for this investigation—M. T. B. and J. J. R.

[3] Von Pechmann, *Ber.*, **15,** 885 (1882).

[4] Kozniewski and Marchlewski, *Bull. Acad. Sci.*, Cracow, 1906, 81.

[5] Kugel, *Ann.*, **299,** 50 (1898).

[6] Rice, *J. Amer. Chem. Soc.*, **45,** 222 (1923); **46,** 214 (1924).

[7] Wolff, *Ann.*, **264,** 252 (1891).